EXTREME-VALUE MOMENT GOODNESS-OF-FIT TESTS

THEODORE P. HILL[1]* AND VICTOR PEREZ-ABREU[2]**

[1] *School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332-0160, U.S.A.*, e-mail: hill@math.gatech.edu

[2] *Centro de Investigacíon en Matemáticas, Guanajuato, Gto 36000, Mexico*, e-mail: pabreu@fractal.cimat.mx

Abstract. A general goodness-of-fit test for scale-parameter families of distributions is introduced, which is based on quotients of expected sample minima. The test is independent of the mean of the distribution, and, in applications to testing for exponentiality of data, compares favorably to other goodness-of-fit tests for exponentiality based on the empirical distribution function, regression methods and correlation statistics. The new minimal-moment method uses ratios of easily-calculated, unbiased, strongly consistent U-statistics, and the general technique can be used to test many standard composite null hypotheses such as exponentiality, normality or uniformity (as well as simple null hypotheses).

Key words and phrases: Goodness-of-fit, scale-parameter families, U-statistics, exponential family, composite exponential hypothesis, minimal moments, extreme values.

1. Introduction

Deciding whether given data comes from a particular class of probability distributions is one of the basic problems of statistics. For example, given a random sample $X_1, X_2, \ldots, X_n$ from an unknown distribution F, it is often desired to test for exponentiality of the data, i.e., to test the null hypothesis

$$H_0 : F \in \{\exp(\beta) : \beta > 0\}, \tag{1.1}$$

(where $\exp(\beta)$ denotes the exponential distribution with density $\beta e^{-\beta x}$ for $x > 0$).

Alternatively, it may be desired to test for uniformity, normality, or other general properties of the underlying distribution, and many goodness-of-fit tests for such composite null hypotheses are available (cf. D'Agostino and Stephens (1986)).

It is the purpose of this article to introduce a general test for *scale-parameter families*, that is, to test the null hypothesis $H_0 : F \in \mathcal{F}$, where $\mathcal{F} = \{F_\beta, \beta > 0\}$ is a family of distributions satisfying $F_\beta(x) = F_1(\beta x)$ for all $\beta > 0$. Such scale-parameter families include the class of exponential distributions in (1.1), as well as many other families such as uniform ($F_\beta \sim U[0, \beta]$, i.e., F_β is uniformly distributed on $[0, \beta]$), and normal ($F_\beta \sim N(0, \beta^2)$). (Note that neither $\{\text{Poisson}(\beta)\}$ nor $\{N(\beta, 1)\}$ are scale-parameter

*Research partially supported by NSF Grant 95-03375.

**Research partially supported by CONACYT Grant 0328PE9506.

families; scale-invariance says that the random variable X has law in the family if and only if βX also does for all $\beta > 0$.)

The test described below is based on the expected value of sample maxima or minima, for which classical U-statistics are standard estimators, and is based on a representation theorem of Hoeffding, on the linearity of expectations of sample extrema, and on the classical U-statistic strong law of large numbers and law of iterated logarithm.

2. Extreme-value moments

Throughout this article, $X, X_1, X_2, \ldots$ are iid random variables with finite mean $E(X)$ and distribution F. The minimal and maximal moments of F are

$$\hat{m}_k(F) = E(X_1 \wedge \cdots \wedge X_k), \qquad \overset{\vee}{m}_k(F) = E(X_1 \vee X_2 \vee \cdots \vee X_k),$$

where k is a positive integer, $a \wedge b = \min\{a,b\}$, $a \vee b = \max\{a,b\}$, and $E(\cdot)$ denotes expectation with respect to the underlying probability. Thus, for example, $\hat{m}_1(F) = \overset{\vee}{m}_1(F) = E(X)$; and $\hat{m}_1(F) = \hat{m}_2(F)$ iff F is degenerate (Dirac point mass). Also note that in contrast to the classical moment $(E(X^k))$ framework, finiteness of $E(X)$ implies that of both $\hat{m}_k$ and $\overset{\vee}{m}_k$ for all k, both $\{\hat{m}_k\}$ and $\{\overset{\vee}{m}_k\}$ are $o(k)$ with decreasing $o(1)$ difference sequences $\{|\hat{m}_k - \hat{m}_{k+1}|\}$, (cf. Hill and Spruill (1994), Lemma 2.2) and both are *linear*

$$\text{(2.1)} \qquad \text{if } G = law(aX+b), \quad \text{then} \quad \hat{m}_k(G) = a\hat{m}_k(F) + b$$
$$\text{and} \quad \overset{\vee}{m}_k(G) = a\overset{\vee}{m}_k(F) + b \text{ for all } k \in \mathbb{N}.$$

A well-known result of Hoeffding (cf. Pollack (1973)) which plays an essential role in this article is that

(2.2) the sequences $\{\hat{m}_k(F)\}$ and $\{\overset{\vee}{m}_k(F)\}$ each determine F (and vice versa).

(Much more is true: even Müntz subsequences $\{k_j : \Sigma 1/k_j = \infty\}$ of $\{\hat{m}_k\}$ and $\{\overset{\vee}{m}_k\}$ (Hill and Spruill (1994)) and moment sequences of *non-extremal* order statistics (Pollack (1973) also determine F.)

Hoeffding's result (2.2) and easy calculations imply, for example, that

$$\text{(2.3)} \qquad \hat{m}_k(F) = \frac{1}{\beta k} \quad \text{for all} \quad k \in \mathbb{N} \text{ iff } F \sim \exp(\beta);$$

and

$$\text{(2.4)} \qquad \overset{\vee}{m}_k(F) = \frac{\beta k}{k+1} \quad \text{for all} \quad k \in \mathbb{N} \text{ iff } F \sim U[0,\beta].$$

(On the other hand, $\{\overset{\vee}{m}_k\}$ for $N(0,1)$ are known in closed form only for $k = 1,2,3,4,5$ (cf. David (1981)).)

There are many statistics which can be used to estimate $\hat{m}_k(F)$ and $\overset{\vee}{m}_k(F)$ from a random sample $X_1, \ldots, X_n$; the standard one for $\hat{m}_k$ is the U-statistic

$$\text{(2.5)} \quad \hat{M}_k(X_1, \ldots, X_n) := \frac{(n-k)!}{n!} \sum \min\{X_{i_1}, \ldots, X_{i_k} : \{i_j\} \text{ distinct}, 1 \le i_j \le n\}$$

which is both an unbiased and (by Hoeffding's (1961) strong law of large numbers for U-statistics) a strongly consistent estimator for $\hat{m}_k$, that is,

$$\hat{M}_k(X_1, \ldots, X_n) \to \hat{m}_k(F) \quad \text{a.s. as } n \to \infty. \tag{2.6}$$

In fact with a second moment hypothesis as well (Var $X < \infty$), central limit and law of iterated logarithm results also hold for convergence of $\hat{M}_k$ to $\hat{m}_k$ (Serfling (1980), p. 191).

We observe that a useful computing formula for (2.5) is

$$\hat{M}_k(X_1, \ldots, X_n) := \frac{1}{\binom{n}{k}} \sum_{i=1}^{n-k+1} \binom{n-i}{k-1} X_{i:n}, \tag{2.7}$$

where $\{X_{i:n}\}$ are the order statistics $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ of $X_1, \ldots, X_n$.

3. Quotient extreme-value moment goodness-of-fit tests

The purpose of this section is to use the linearity of the minimal moments, linearity condition (2.1), Hoeffding's representation property of the extremal moments (2.2), and the U-statistic strong law of large numbers (2.6) to introduce a goodness-of-fit test for scale-parameter families which is based on quotients of the minimal moments and U-statistics. For simplicity only positive variables and the minimal moments $\{\hat{m}_k\}$ will be addressed, and the extension to general variables and to the maximal moments $\{\check{m}_k\}$ is left to the interested reader.

For each $k \in \mathbb{N}$, let $q_k(F) := \hat{m}_k(F)/\hat{m}_{k+1}(F)$ and

$$q_k(X_1, \ldots, X_n) := \hat{M}_k(X_1, \ldots, X_n)/\hat{M}_{k+1}(X_1, \ldots, X_n). \tag{3.1}$$

The next theorem gives a minimal-moment characterization of scale-parameter families which is the basis for the test statistics $Q_{p,F}$ defined below.

THEOREM 3.1. *Suppose $X_1, X_2, \ldots$ are iid with distribution F, and $\mathcal{F} = \{F_\beta\}$ is a scale-parameter family of (integrable) random variables. Then the following are equivalent:*

(i) *$F = F_\beta$ for some $\beta > 0$;*
(ii) *$q_k(F) = q_k(F_1)$ for all $k \in \mathbb{N}$;*
(iii) *$q_k(X_1, \ldots, X_n) \to q_k(F_1)$ a.s. as $n \to \infty$, for all $k \in \mathbb{N}$.*

PROOF. "(i) $\Rightarrow$ (ii)." Suppose $F = F_\beta$ for some $\beta > 0$. Then by (2.1), $q_k(F) = q_k(F_\beta) = \beta\hat{m}_k(F_1)/\beta\hat{m}_{k+1}(F_1) = q_k(F_1)$ for all $k \in \mathbb{N}$.

"(ii) $\Rightarrow$ (i)." Suppose $q_k(F) = q_k(F_1) =: q_k$ for all $k \in \mathbb{N}$, and let $\mu = \hat{m}_1(F) > 0$ and $\lambda = \hat{m}_1(F_1) > 0$. Then by the definition of $\{q_k\}$, $\hat{m}_{k+1}(F) = \mu/\prod_{j=1}^k q_j$ and $\hat{m}_{k+1}(F_1) = \lambda/\prod_{j=1}^k q_j$, so $\hat{m}_k(F) = \frac{\mu}{\lambda}\hat{m}_k(F_1)$ for all $k \in \mathbb{N}$, which by linearity (2.1) and Hoeffding's representation theorem (2.2) implies that F is in the scale-parameter family containing F_1, so $F = F_\beta$ for some $\beta > 0$.

"(i) $\Rightarrow$ (iii)." Suppose $F = F_\beta$ for some $\beta > 0$, and fix $k \in \mathbb{N}$. By the U-statistic strong law of large numbers (2.6), $q_k(X_1, \ldots, X_n) \to q_k(F_\beta)$ a.s. $n \to \infty$. But by linearity (2.1) and the definition of $\{q_k\}$, $q_k(F_1) = q_k(F_\beta)$.

"(iii) $\Rightarrow$ (ii)." Immediate by (2.6) and the definition of $\{q_k\}$. □

COROLLARY 3.2. (Exponential family case). *Let $X_1, X_2, \ldots$ be iid with finite positive mean and with distribution F. Then the following are equivalent*:

(i) $F \sim \exp(\beta)$ *for some* $\beta > 0$;

(ii) $q_k(F) = (k+1)/k$ *for all* $k \in \mathbb{N}$;

(iii) $q_k(X_1, \ldots, X_n) \to (k+1)/k$ *a.s. for all* $k \in \mathbb{N}$.

PROOF. Immediate from Theorem 3.1 and the $\{\hat{m}_k\}$ for exponentials in (2.3). □

Test statistics $\{Q_{p,F} : 1 \le p \le \infty\}$ will now be defined which measure the discrepancy between the sample quotients (of minimal moments) and the true quotients of the scale-parameter family containing F. The distances used are the standard ℓ_p distances, and the quotients are simply those of the classical U-statistics estimators (2.5) of the minimal moments $\{\hat{m}_k(F)\}$. Here $\|(x_1, \ldots, x_m)\|_p$ denotes the usual ℓ_p norm $\left(\sum_{i=1}^m |x_i|^p\right)^{1/p}$ for $1 \le p < \infty$, and the sup norm $\max\{|x_1|, \ldots, |x_m|\}$ for $p = \infty$. The composite null hypothesis for the random sample to be from the given scale-parameter family (e.g. (1.1)) will then be rejected iff the discrepancy $Q_{p,F}$ is sufficiently large.

Let $\{a_n\}$ be an unbounded increasing sequence of positive integers with $a_n = o((n/\log\log n)^{1/2})$.

DEFINITION 3.3. Let $\mathcal{F} = \{F_\beta\}$ be a scale-parameter family, and let $X_1, X_2, \ldots, X_n$ be iid with finite positive mean. For each p, $1 \le p \le \infty$, let

$$Q_{p,F_1}(X_1, \ldots, X_n) := \|\vec{q}_n(X_1, \ldots, X_n) - \vec{q}_n(F_1)\|_p, \tag{3.2}$$

where $\vec{q}_n(X_1, \ldots, X_n) = (q_1(X_1, \ldots, X_n), \ldots, q_{a_n}(X_1, \ldots, X_n))$ and $\vec{q}_n(F_1) = (q_1(F_1), \ldots, q_{a_n}(F_1))$.

Thus the statistic (3.2) of a random sample of size n is the ℓ_p distance between the first a_n quotients of minimal moments of the sample, and those of the distribution. Thus for $a_n = \lceil n^{1/3} \rceil + 6$ (as was used in the simulation below), if $n = 20$ the first 9 minimal-moment quotients are compared, and if $n = 50$ the first 10 are compared. (Here $\lceil x \rceil = \min\{i \in \mathbb{N} : i \ge x\}$.)

THEOREM 3.4. *Let $X_1, X_2, \ldots$ be positive iid random variables with distribution F, let $\mathcal{F} = \{F_\beta\}$ be a scale-parameter family, and let $1 \le p \le \infty$. If $Q_{p,F_1}(X_1, \ldots, X_n) \to 0$ a.s., then $F \in \mathcal{F}$. Conversely, if F_1 has finite variance, then for all $F \in \mathcal{F}$, $Q_{p,F_1}(X_1, \ldots, X_n) \to 0$ a.s.*

PROOF. Fix p, $1 \le p \le \infty$.

"$\Rightarrow$" By way of contradiction, suppose $F \notin \mathcal{F}$. Then by Theorem 3.1 there exists an $m \in \mathbb{N}$, and $d > 0$ so that $|q_m(F) - q_m(F_1)| = d$. By (2.6) and the definition of $\{q_m\}$, $q_m(X_1, \ldots, X_n) \to q_m(F)$ a.s. as $n \to \infty$, so $\liminf_{n\to\infty} Q_{p,F_1}(X_1, \ldots, X_n) \ge |q_m(F) - q_m(F_1)| = d > 0$ a.s.

Conversely, suppose that F_1 has finite variance σ^2. Since the $\{q_k\}$ are scale invariant, assume without loss of generality that $F = F_1$. If $\sigma^2 = 0$, then $\mathcal{F}$ is the class of (positive) constant (Dirac measure) distributions, and trivially $q_k(X_1, \ldots, X_n) = q_k(F_\beta) = 1$ a.s. for all k and β, so $Q_{p,F_1}(X_1, \ldots, X_n) \equiv 0$ a.s. If $\sigma^2 > 0$, then since $0 < \mathrm{Var}(X_1) < \infty$ if and only if $0 < \mathrm{Var}(\min\{X_1, \ldots, X_k\}) < \infty$, it follows from Serfling's law of iterated logarithm for U-statistics (Serfling (1980), Theorem C, p. 191), using kernel $h(x_1,\dots,x_k)=\min\{x_1,\dots,x_k\}$, that

$$\frac{\hat{M}_k(X_1,\dots,X_n)-\hat{m}_k(F_1)}{\sqrt{\dfrac{\log\log n}{n}}}=O(1) \qquad \text{a.s.}$$

Hence

$$\begin{aligned} q_k(X_1,\dots,X_n) &= \frac{\hat{M}_k(X_1,\dots,X_n)}{\hat{M}_{k+1}(X_1,\dots,X_n)} \\ &= \frac{\hat{m}_k(F_1)+O(\sqrt{(\log\log n)/n})}{\hat{m}_{k+1}(F_1)+O(\sqrt{(\log\log n)/n})} \\ &= \frac{\hat{m}_k(F_1)+O(\sqrt{(\log\log n)/n})}{\hat{m}_{k+1}(F_1)} \qquad \text{a.s.} \end{aligned}$$

Since $\hat{m}_{k+1}(F_1)>0$, this implies that

$$|q_k(X_1,\dots,X_n)-q_k(F_1)|=O(\sqrt{(\log\log n)/n}) \qquad \text{a.s.} \tag{3.3}$$

Since

$$\|\vec{q}_n(X_1,\dots,X_n)-\vec{q}_n(F_1)\|_p\le\sum_{k=1}^{a_n}|q_k(X_1,\dots,X_n)-q_k(F_1)|,$$

it follows from (3.3) that

$$Q_{p,F_1}(X_1,\dots,X_n)=O(a_n\sqrt{(\log\log n)/n}) \qquad \text{a.s.,}$$

and the conclusion then follows since $a_n=o((n/\log\log n)^{1/2})$. □

COROLLARY 3.5. (Exponential family case). *Let $X_1,X_2,\dots$ be positive iid with distribution F, and let $\mathcal{F}=\{\exp(\beta):\beta>0\}$. Then for all $1\le p\le\infty$,*

$$\left(\sum_{k=1}^{\lceil n^{1/3}\rceil}|q_k(X_1,\dots,X_n)-(k+1)/k|^p\right)^{1/p}\to 0 \textit{ a.s. iff } F \textit{ is } \exp(\beta) \textit{ for some } \beta>0.$$

PROOF. Immediate from Theorem 3.4 and (2.3). □

Analogs of Corollary 3.5 to other standard scale-parameter families are also easy to construct. In case the minimal or maximal moments $\{\hat{m}_k\}$ and $\{\check{m}_k\}$ are not available in closed form (as in the Gaussian case), they may easily be approximated by Monte Carlo simulation using the U-statistic (2.5).

4. Application to tests for exponentiality

In this section, application of the above ideas is given for testing whether data comes from an exponential distribution $\exp(\beta)$ for some $\beta>0$; that is, for testing the null hypothesis in (1.1). Although there are various other tests for exponentiality, such as the Pearson chi-square test and modern adaptations, empirical distribution function tests and regression tests, and tests based on the classical sample moments (cf. D'Agostino and Stephens (1986), Chapter 10, Ascher (1990), and Baringhaus and Henze (1991)), the new $Q_{p,F}$ test is versatile (through choice of p and numbers of moment-quotients compared), is easy to implement, has limiting distribution which is parameter free, and is based on classical U-statistics about which a great deal is known. Since Poisson processes are characterized by exponential interarrival times, these tests may also be used for testing whether given point process data comes from a homogeneous Poisson process. Analogs of the tables in this section to $Q_{p,F}$ tests for uniformity, normality and other scale-parameter families are calculated similarly.

Table 1. Critical points for $Q_{p,F}$ test for exponentiality.

	$n=10$			$n=20$		
	$\alpha=.001$	$\alpha=.01$	$\alpha=.05$	$\alpha=.001$	$\alpha=.01$	$\alpha=.05$
$Q_{1,F}$	58.944	12.175	5.150	3.192	2.285	1.623
$Q_{2,F}$	55.219	9.037	2.557	1.421	1.017	0.704
$Q_{\infty,F}$	55.195	8.981	2.185	1.307	0.875	0.557
	$n=30$			$n=40$		
$Q_{1,F}$	2.400	1.741	1.278	1.919	1.397	1.058
$Q_{2,F}$	1.058	0.768	0.540	0.895	0.632	0.449
$Q_{\infty,F}$	0.997	0.648	0.442	0.829	0.552	0.370
	$n=50$			$n=100$		
$Q_{1,F}$	1.551	1.243	0.914	1.125	0.873	0.648
$Q_{2,F}$	0.767	0.550	0.393	0.485	0.367	0.276
$Q_{\infty,F}$	0.663	0.479	0.330	0.430	0.318	0.232

The critical points above for different values of α, p and n were calculated by simulation: n exponential observations, $X_1, \ldots, X_n$ were generated and the statistic $Q_{p,F}$ computed (using (3.2) and Corollary 3.2 (i)–(ii)); this was repeated for a total of 10,000 iterations, and the corresponding empirical α-quantiles were determined

Comparisons of the $Q_{p,F}$ test with other standard tests for exponentiality are given in the following table against standard alternatives to the exponential distribution including types of increasing, decreasing, and non-monotone failure rates (recall the exponential has constant failure rate). The alternatives used (cf. D'Agostino and Stephens (1986), p. 452–453) are the IFR types χ_4^2, $U[0,1]$, Weib(1.5), $|N|$, Gamma(1.4), Gamma(2.0), Beta(2,1) and Beta(1,2), DFR types χ_1^2, Weib(0.8), $|C|$, Gamma(0.5), and Gamma(0.7), and the non-monotone failure rate types LN(1), LN(1.5) and Beta(0.5,1) where:

Weib(γ) is the Weibull distribution with density, $f(x) = \gamma x^{\gamma-1}\exp(-x^\gamma)$, $x > 0$;
$|N|$ is the distribution of $Y = |N|$ where N is $N(0,1)$;
LN refers to the lognormal density $f(x) = \text{const} \cdot x^{-1} \cdot \exp(-(\log x)^2/2)$, $x > 0$;
$|C|$ is the distribution of $Y = |C|$ where C is standard Cauchy;
exp(1) is mean-1 exponential distribution.

Standard tests against which the $Q_{p,F}$ test is compared in Table 2 include the Anderson-Darling test A^2, the Cramer-von Mises test W^2, Stephens' W_S regression test, Baringhaus and Henze's T_1 test, Moran's test M, Patwardhan'(s) $Q(1)$ test, Gnedenko's F-test $Q(R)$, and the Cox-Oakes test CO, details of which may be found in D'Agostino and Stephens (1986), Ascher (1990), and Baringhaus and Henze (1991).

Table 2A. Estimated powers, in percent, of $Q_{p,F}$ test for exponentiality at 5% level.

$n = 20$; 5000 trials											
	$Q_{1,F}$	$Q_{2,F}$	$Q_{\infty,F}$	A^2	W^2	W_S	T_1	M	$Q(1)$	$Q(R)$	CO
Null Hypothesis											
exp(1)	6	5	4	4	4	4	4	5	4	5	5
Alternatives											
Increasing Failure Rate											
χ_4^2	9	7	10	45	47	35	52	56	15	26	57
U[0,1]	7	12	28	63	66	69	60	46	27	60	59
Weib(1.5)	8	7	11	45	47	40	51	51	15	30	56
$\|N\|$	2	2	3	17	21	17	20	18	9	17	22
Gamma(1.4)	2	1	2	13	15	9	15	17	6	9	17
Gamma(2.0)	10	8	10	45	48	34	53	57	16	25	57
Beta(2,1)	75	94	99	100	100	100	100	100	87	99	100
Beta(1,2)	3	2	4	22	25	25	24	20	11	24	25
Decreasing Failure Rate											
χ_1^2	79	78	65	70	51	37	60	75	26	46	71
Weib(0.8)	34	36	34	27	20	19	25	28	12	20	26
$\|C\|$	57	68	71	63	63	68	64	53	53	57	57
Gamma(0.5)	79	78	67	71	52	36	61	77	25	47	72
Gamma(0.7)	38	36	31	27	18	15	23	31	11	19	27
Non-monotone Failure Rate											
LN(1)	4	8	10	12	13	17	11	8	10	5	9
LN(1.5)	5	9	11	14	15	17	13	10	11	7	10
Beta(0.5,1)	52	38	12	40	18	3	13	39	22	14	28

From our power study we conclude that for large sample sizes the proposed tests $Q_{p,F}$ behave well against a large selection of alternatives and are competitive with other well known goodness of fit tests for exponentiality. For small samples, the statistics $Q_{p,F}$ (especially $Q_{2,F}$) behave very well against DFR alternatives, but poorly against IFR distributions. It is expected that there will be the opposite situation if maximal moments are used instead of minimal moments.

Remark. Corollary 3.2 ((i)$\Leftrightarrow$(ii)) suggests an alternative test for exponentiality: estimate $\hat{m}_k$ with the U-statistics $\hat{M}_k$ in (2.5), and then test $\hat{M}_k^{-1}$ for *linearity* in k. A similar linearity-based goodness-of-fit test for Poisson data was obtained by Nakamura and Perez-Abreu (1993).

5. Remarks

a) The same ideas can be used to test for simple null hypotheses by just comparing the sample minimal moments with true minimal moments. For example, the null hypothesis $H_0 : F = \exp(1)$ should be rejected iff the sample minimal moments $\{\hat{M}_k(X_1, \ldots, X_n)\}$ are not close to the true minimal moments $\{1/k\}$ (cf. (2.3)). Although the above statistics $Q_{p,F}$ were designed to test for scale-parameter families only,

Table 2B. Estimated powers, in percent, of $Q_{p,F}$ test for exponentiality at 5% level.

$n = 50$; 5000 trials											
	$Q_{1,F}$	$Q_{2,F}$	$Q_{\infty,F}$	A^2	W^2	W_S	T_1	M	$Q(1)$	$Q(R)$	CO
Null Hypothesis											
exp(1)	5	5	5	5	5	5	5	5	5	5	5
Alternatives											
Increasing Failure Rate											
χ_4^2	84	85	78	92	90	77	94	96	40	64	96
U[0,1]	50	91	97	99	98	100	96	79	53	95	94
Weib(1.5)	71	83	82	91	90	86	94	93	35	73	95
$\|N\|$	16	29	36	45	49	52	50	36	14	43	49
Gamma(1.4)	21	20	18	32	32	23	37	40	9	19	41
Gamma(2.0)	85	84	78	91	90	77	94	96	41	64	95
Beta(2,1)	100	100	100	100	100	100	100	100	100	100	100
Beta(1,2)	16	37	49	57	59	75	59	40	19	54	58
Decreasing Failure Rate											
χ_1^2	98	98	93	96	90	63	94	98	40	79	97
Weib(0.8)	59	62	57	52	43	31	51	52	16	38	52
$\|C\|$	89	95	95	93	93	94	93	86	83	89	91
Gamma(0.5)	98	98	93	96	90	63	94	98	39	79	97
Gamma(0.7)	63	61	50	52	38	21	46	57	13	31	55
Non-monotone Failure Rate											
LN(1)	30	18	16	34	29	28	17	14	23	8	11
LN(1.5)	30	19	17	35	30	28	18	15	22	10	13
Beta(0.5,1)	79	64	15	76	48	5	24	63	35	15	46

a similar statistic can test for (finite-moment) families in which both scale and location are unknown, since by linearity (2.1), $(\hat{m}_{k+1} - \hat{m}_k)/(\hat{m}_{k+2} - \hat{m}_{k+1})$ is independent of both scale *and* location.

b) The statistics $Q_{p,F}$ described above give equal weight to proximity to all minimal moments, but perhaps nonuniform weights (e.g., more weight on proximity to means than higher minimal moments) will lead to tests which are more powerful against standard alternatives; this has not yet been studied by the authors.

c) Although much is known about the distribution of sample extrema (cf. Resnick (1987)) and about convergence and limiting distributions of the U-statistics $\{\hat{M}_k\}$, the approximate and limiting distributions of $Q_{p,F}$ or of the ratios $\hat{M}_k/\hat{M}_{k+1}$ are not known to the authors.

d) The statistics $Q_{p,F}$ have several additional advantages. First, in many applications such as those involving time to failure, the sample minima (and their quotients, which measure the advantage of adding an additional component to the system) are natural objects of independent interest.

Second, they are (via Theorem 3.4) one of the only goodness-of-fit tests for composite scale-parameter null hypotheses which are consistent against *any* finite-moment alternative to the null hypothesis. And third, since they depend only on existence of a finite first moment, they may prove useful in testing distributions such as stable laws which do not have finite variance.

e) As pointed out to us by Michael Stephens, the quotient test described above could be useful in the presence of censoring, since the minimum, for example, is insensitive to deletion of high values of the sample. This is probably also the reason for the better performance of the $Q_{p,F}$ statistics against DFR alternatives than against IFR alternatives. In addition, an even more censoring-robust test could be easily designed using a statistic based on $q_c, \ldots, q_{c+d}$ instead of $q_1, \ldots, q_{d+1}$, since the moments $\{\hat{m}_k, k \geq c\}$, also determine the distribution uniquely.

f) Since the U-statistic strong law of large numbers also holds for large classes of dependent random variables such as ergodic stationary sequences (Aaronson, *et al.* (1996)), the same $Q_{p,F}$ statistical tests can be used for many applications in which the data $X_1, X_2, \ldots$ are not necessarily independent.

Acknowledgements

The first author wishes to thank CIMAT in Guanajuato and the Free University of Amsterdam for their hospitality during visits in which some of this research was done. The authors wish to thank Lisa Bloomer and Armando Dominguez for their assistance in running the simulations used to obtain the data in Tables 1 and 2, and to thank the Associate Editor, the referees and Annette Rohrs for a number of useful suggestions.

References

Aaronson, J., Burton, R., Dehling, H., Gilat, D., Hill, T. and Weiss, B. (1996). Strong laws for L- and U-statistics, *Trans. Amer. Math. Soc.*, **348**, 2845–2866.

Ascher, S. (1990). A survey of tests for exponentiality, *Comm. Statist. Theory Methods*, **19**, 1811–1825.

Baringhaus, L. and Henze, N. (1991). A class of consistent tests for exponentiality based on the empirical Laplace transform, *Ann. Inst. Statist. Math.*, **43**, 551–564.

D'Agostino, R. and Stephens, M. (1986). *Goodness of Fit Techniques*, Marcel Dekker, New York.

David, H. (1981). *Order Statistics*, 2nd ed., Wiley, New York.

Hill, T. and Spruill, M. (1994). On the relationship between convergence in distribution and convergence of expected extremes, *Proc. Amer. Math. Soc.*, **121**, 1235–1243.

Hoeffding, W. (1961). The strong law of large numbers for U-statistics, Mimeo Series **302**, Institute of Statistics, University of North Carolina.

Nakamura, M. and Perez-Abreu, V. (1993). Use of an empirical probability generating function for testing a Poisson model, *Canad. J. Statist.*, **21**, 149–156.

Pollack, M. (1973). On equal distributions, *Ann. Statist.*, **1**, 180–182.

Resnick, S. (1987). *Extreme Values, Regular Variation and Point Processes*, Springer, New York.

Serfling, R. (1980). *Approximation Theorems of Mathematical Statistics*, Wiley, New York.